Optimal Uncertainty Quantification

Bounds, Predictions and Experimental Design

Tim Sullivan

tjs@caltech.edu

California Institute of Technology

http://www.psaap.caltech.edu/

SciDAC 2011 — Uncertainty Quantification Session

Denver, Colorado, U.S.A.

14 July 2011

CALTECH PSAAP

National Nuclear Security Administration

Collaborators — Caltech PSAAP Center

Experimental

A. Rosakis
G. Ravichandran
M. Adams
J. Mihaly
J. Brown
L. Bodelot
A. Kidane
K. John

Software

M. Aivazis
M. Stalzer
M. McKerns
S. Brunett
J. Cummings
J. Lindheim
S. Lombeyda
B. Li
L. Strand

Management

M. Ortiz
M. Stalzer

UQ

H. Owhadi
T. J. Sullivan
M. McKerns
B. Li
C. Scovel (LANL)

Fluids

D. Meiron
D. Pullin
P. Barton
J. Cummings
G. Ward
A. L. Ortega

Solids

M. Ortiz
K. Bhattacharya
W. A. Goddard III
A. van de Walle
A. Pandolfi
A. Jaramillo-Botero
J. Amelang
L. Djodom Fokuoa
A. Kowalski
X. Wang
A. Richards
L. Miljacic
B. Li
A. Bompadre
S. Demers
P. Theofanis
J. Tahir-Kheli
P. Cesana
L. Perotti
D. Kochmann

Outline

Portions of this work were supported by the United States Department of Energy National Nuclear Security Administration under Award Number DE-FC52-08NA28613 through the California Institute of Technology's ASC/PSAAP Center for the Predictive Modeling and Simulation of High Energy Density Dynamic Response of Materials.

Hypervelocity Impact

- **PSAAP:** Advance the prediction of the behaviour of complex systems with quantified margins and uncertainties.
- **Caltech Center:** Develop predictive science methods focusing on high-energy-density dynamic response of materials.
- Overarching ASC-class application is hypervelocity impact:

$$\left.\begin{array}{c}\textbf{impactor}\\ \textit{1–20 mg}\\ \hline \text{steel}\\ \text{nylon}\\ \text{tantalum}\end{array}\right\} \xrightarrow[\text{normal/oblique}]{2\text{–}10\,\mathrm{km/s}} \left\{\begin{array}{c}\textbf{target plate}\\ \textit{1–5 mm}\\ \hline \text{steel}\\ \text{aluminium}\\ \text{tantalum}\end{array}\right.$$

Figure: Impact flash from a 7.9 km/s hypervelocity test. (NASA Ames Research Center)

Why Hypervelocity Impact?

- Generation of states of matter of interest:
  - high pressures (160–800 GPa);
  - high strain rates (up to 1,000 s^{-1});
  - high temperatures (4,000–36,000 K).
- Multiphysics, complex material behaviour:
  - melting, vaporization, ionization, plasma;
  - luminescence and radiative transport;
  - hydro instabilities, mixed-phase flows, mixing;
  - solid-solid phase transitions, high-strain-rate deformation, thermo-mechanical coupling;
  - fracture, fragmentation, spall and ejecta, deformation instabilities *e.g.* shear banding.

$\Rightarrow$ Experimental and modelling challenges.

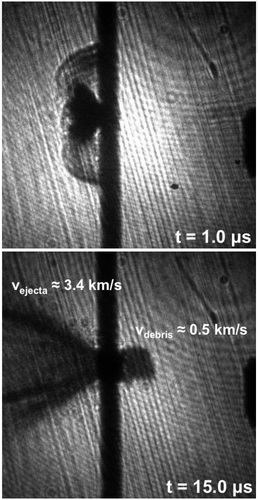
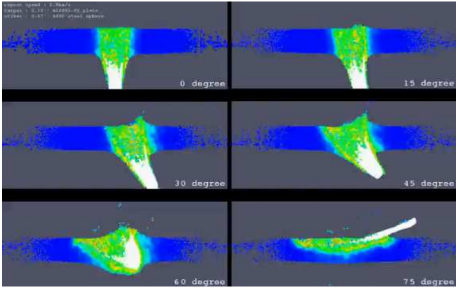


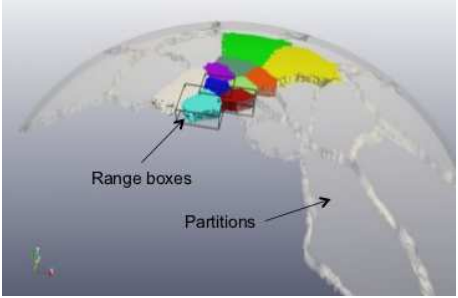
Figure: 25 ns laser side-lit exposures of a 5.4 km/s nylon-on-Al impact in the Caltech SPHIR facility.

Parallel Optimal Transportation Mesh-Free Solver

- Optimal Transport
  incremental variational update formulation with geometrically exact discrete Lagrangians
- Mesh-Free
  physical data carried by material points, with mesh-free local max-ent nodal interpolation
- Parallelization
  asynchronous shadow nodes in overlapping range boxes
  - linear scaling to 256 cores
  - good scaling to 2048 cores

  (bound by communications cost)





Uncertainty Quantification

Or: How I Learned to Stop Worrying and Bound Things

Optimization-Driven UQ

Bounds Mean Optimizations!

- Conventional worst/best-case design is an optimization problem over possible design and operation parameters:

$$\min_{x \in \mathcal{X}} G(x), \qquad \max_{x \in \mathcal{X}} G(x).$$

- Insufficient to make statements about *e.g.* probabilities of events.
- We want to handle generic information about the probability distributions and response functions, which are in general incompletely specified.

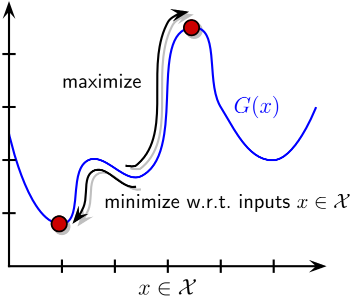


Figure: Optimizing $G(x)$ over $x \in \mathcal{X}$ yields deterministic worst- and best-case outcomes. What if the distribution of the inputs is only *partially* known? (*I.e.* non-parametric epistemic uncertainty.)

Optimization-Driven UQ

Bounds Mean Optimizations!

- Conventional worst/best-case design is an optimization problem over possible design and operation parameters:

$$\min_{x \in \mathcal{X}} G(x), \qquad \max_{x \in \mathcal{X}} G(x).$$

- Insufficient to make statements about *e.g.* probabilities of events.
- We want to handle generic information about the probability distributions and response functions, which are in general incompletely specified.

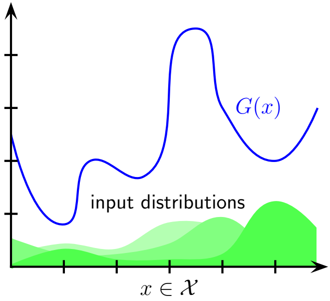


Figure: Optimizing $G(x)$ over $x \in \mathcal{X}$ yields deterministic worst- and best-case outcomes. What if the distribution of the inputs is only *partially* known? (*I.e.* non-parametric epistemic uncertainty.)

Optimization-Driven UQ

Bounds Mean Optimizations!

- Conventional worst/best-case design is an optimization problem over possible design and operation parameters:

$$\min_{x \in \mathcal{X}} G(x), \qquad \max_{x \in \mathcal{X}} G(x).$$

- Insufficient to make statements about *e.g.* probabilities of events.
- We want to handle generic information about the probability distributions and response functions, which are in general incompletely specified.

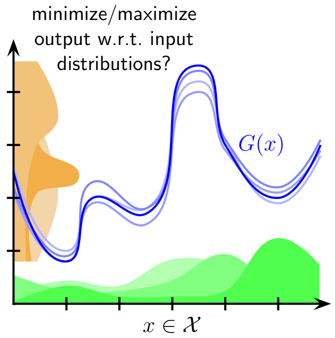


Figure: Optimizing $G(x)$ over $x \in \mathcal{X}$ yields deterministic worst- and best-case outcomes. What if the distribution of the inputs is only *partially* known? (*I.e.* non-parametric epistemic uncertainty.)

Optimal Uncertainty Quantification (OUQ)

- OUQ is a mathematically rigorous formulation of UQ that places information at the centre of the problem — items of information are viewed as constraints.
- Particularly suited the regime of high-consequence decision-making with incomplete information.
- Naturally generalizes classical interval analysis and optimization-based UQ methods to the probabilistic regime.
- Basic idea: pick a quantity of interest and optimize (minimize/maximize) with respect to the scenarios compatible with your current state of knowledge.

UQ Problems

- Reliability
- Certification
- Verification
- Validation
- Extrapolation
- Prediction
- Sensitivity
- Model Reduction
- ...

Owhadi *& al.* (2010)

http://arxiv.org/abs/1009.0679

OUQ Paradigm

- Abstract system $G \colon \mathcal{X} \to \mathcal{Y}$ with random inputs X with probability distribution $\mathbb{P} \in \mathcal{P}(\mathcal{X})$ — but the pair $(G, \mathbb{P})$ is imperfectly known!
- Quantity of interest $\mathbb{E}[q_G]$, *e.g.* the mean $\mathbb{E}[G]$, or the probability of failure $\mathbb{P}[G \in \mathcal{F}] \equiv \mathbb{E}[\mathbb{1}[G \in \mathcal{F}]]$ for some critical/failure event $\mathcal{F}$.
- Feasible set of admissible scenarios that could be the reality $(G, \mathbb{P})$:

$$\mathcal{A} := \left\{ (g, p) \,\middle|\, \begin{array}{c} (g \colon \mathcal{X} \to \mathcal{Y}, p \in \mathcal{P}(\mathcal{X})) \text{ is consistent with} \\ \text{all given information about the real system } (G, \mathbb{P}) \\ (\textit{e.g.} \text{ legacy data, first principles, expert judgement}) \end{array} \right\}.$$

- Optimal bounds on $\mathbb{E}[q_G]$ found by minimizing/maximizing $\mathbb{E}_p[q_g]$ over $(g, p) \in \mathcal{A}$:

$$\min q \leq \min_{(g,p) \in \mathcal{A}} \mathbb{E}_p[q_g] \leq \mathbb{E}[q_G] \leq \max_{(g,p) \in \mathcal{A}} \mathbb{E}_p[q_g] \leq \max q.$$

Reduction of OUQ Problems — LP Analogy

Dimensional Reduction

- *A priori*, OUQ problems are infinite-dimensional, non-convex, highly-constrained, global optimization problems.
- However, they can be reduced to equivalent finite-dimensional problems in which the optimization is over the extremal scenarios of $\mathcal{A}$.
- The dimension of the reduced problem is proportional to the number of probabilistic inequalities that describe $\mathcal{A}$.

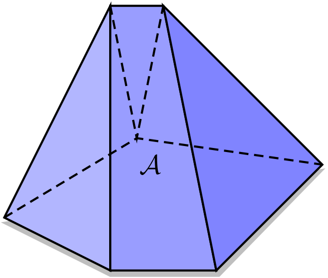


Figure: Just as a linear program finds its extreme value at the extremal points of a convex domain in $\mathbb{R}^n$, OUQ problems reduce to searches over finite-dimensional families of extremal scenarios.

Reduction of OUQ Problems — LP Analogy

Dimensional Reduction

- *A priori*, OUQ problems are infinite-dimensional, non-convex, highly-constrained, global optimization problems.
- However, they can be reduced to equivalent finite-dimensional problems in which the optimization is over the extremal scenarios of $\mathcal{A}$.
- The dimension of the reduced problem is proportional to the number of probabilistic inequalities that describe $\mathcal{A}$.

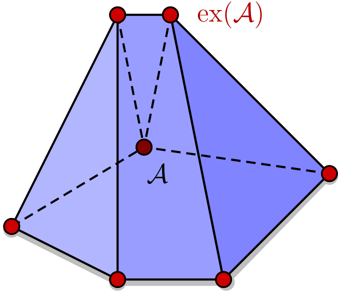


Figure: Just as a linear program finds its extreme value at the extremal points of a convex domain in $\mathbb{R}^n$, OUQ problems reduce to searches over finite-dimensional families of extremal scenarios.

Reduction of OUQ Problems — Example

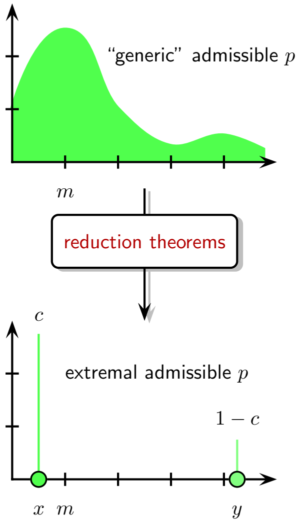
Example

If we are interested in bounding $\mathbb{P}[X \geq a]$ where X is a random variable known to satisfy

$$X \geq 0 \qquad \text{and} \qquad \mathbb{E}[X] = m$$

then we find the extreme values by searching among probability distributions that are just two point masses, *i.e.* of the form

$$\begin{aligned} p &= c\delta_x + (1-c)\delta_y \\ \text{subject to: } & x, y \geq 0 \\ & 0 \leq c \leq 1 \\ & m = cx + (1-c)y. \end{aligned}$$



(Non-)Propagation of Information

One can consider hierarchies (directed acyclic graphs) of OUQ modules:

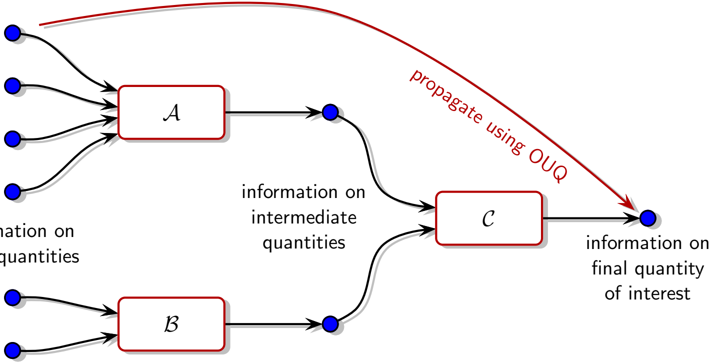


Figure: Because OUQ is a *sharp* information propagation scheme, the results of sensitivity analysis ("inverse OUQ") give non-trivial insights into the roles of the various pieces of input information. Some inputs may even be irrelevant!

(Non-)Propagation of Information

One can consider hierarchies (directed acyclic graphs) of OUQ modules:

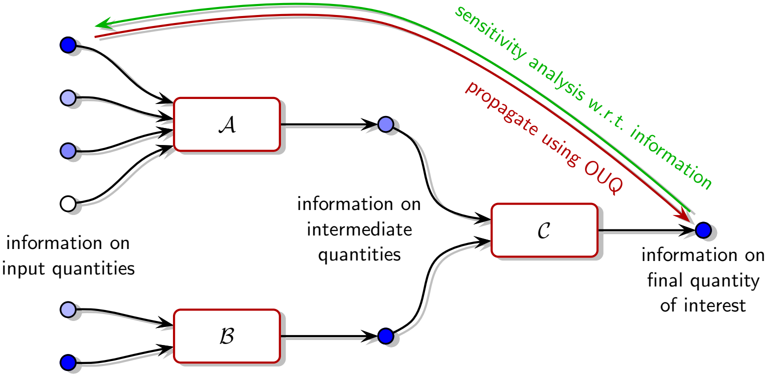


Figure: Because OUQ is a *sharp* information propagation scheme, the results of sensitivity analysis ("inverse OUQ") give non-trivial insights into the roles of the various pieces of input information. Some inputs may even be irrelevant!

OUQ-Driven Experimental Planning

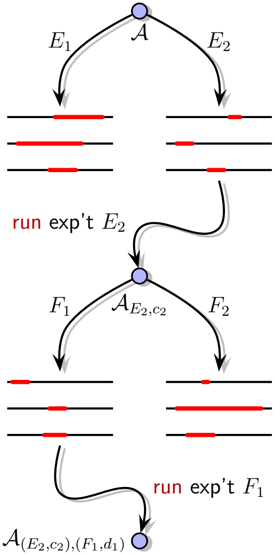
- Range of prediction for q given $\mathcal{A}$:

$$\mathcal{R}(q|\mathcal{A}) := \max_{(g,p)\in\mathcal{A}} \mathbb{E}_p[q_g] - \min_{(g,p)\in\mathcal{A}} \mathbb{E}_p[q_g],$$

$$\mathcal{R}(q|\mathcal{A}) \text{ small} \longleftrightarrow \mathcal{A} \text{ very predictive.}$$

- Let $\mathcal{A}_{E,c}$ denote those scenarios in $\mathcal{A}$ that are consistent with getting outcome c from some experiment E.
- The optimal next experiment E^* satisfies a minimax criterion, *i.e.* E^* is the most predictive even in its least predictive outcome:

$$E^* \text{ minimizes } E \mapsto \max_{\substack{\text{outcomes } c\\ \text{of } E}} \mathcal{R}(q|\mathcal{A}_{E,c}).$$

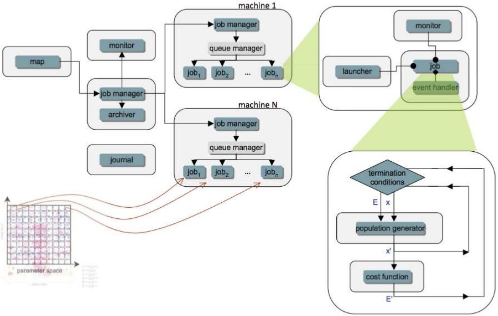


The mystic Optimization Framework

OUQ has developed in symbiosis with the mystic optimization framework.

mystic

- open-source Python
- simple interface to massively parallel optimization
- seamless use of heterogeneous resources
- OUQ calculations with hundreds of variables
- pre-applied constraints
- swappable optimizers launched as services



McKerns *& al.* (2010)

http://dev.danse.us/trac/mystic

References

- Optimal UQ methods
  - H. Owhadi, C. Scovel, T. J. Sullivan, M. McKerns & M. Ortiz. "Optimal Uncertainty Quantification" Submitted to *SIAM Review*. Preprint: http://arxiv.org/abs/1009.0679
  - D. Meyer, M. McKerns, M. Ortiz, H. Owhadi, T. J. Sullivan & F. Theil. "Optimal Uncertainty Quantification for Legacy Data Observations of Lipschitz Functions" In Preparation.
- `mystic` framework
  - http://dev.danse.us/trac/mystic
  - M. McKerns, L. Strand, T. J. Sullivan, A. Fang & M. Aivazis. "Building a Framework for Predictive Science" Submitted to *Proceedings of SciPy 2011*.
- Optimal Transportation Mesh-Free methods
  - B. Li, F. Habbal & M. Ortiz. "Optimal transportation meshfree approximation schemes for fluid and plastic flows." *Int. J. Numer. Meth. Engrg.* **83**(12):1541–1579 (2010). doi:10.1002/nme.2869